Docket No. NEB-181

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
APPLICATION FOR UNITED STATES LETTERS PATENT

INVENTOR (S): Romualdas Vaisvila
Richard D. Morgan
Rebecca B. Kucera
Toby E. Claus
Elisabeth A. Raleigh

TITLE: METHOD FOR CLONING AND PRODUCING THE *Mse*I RESTRICTION ENDONUCLEASE

ATTORNEY: Gregory D. Williams
General Counsel
NEW ENGLAND BIOLABS INC.,
32 Tozer Road
Beverly, Massachusetts 01915
(978) 927-5054; Ext. 292

EXPRESS MAILING LABEL NO.: EK589681297US

METHOD FOR CLONING AND PRODUCING THE MseI RESTRICTION ENDONUCLEASE

BACKGROUND OF THE INVENTION

Restriction endonucleases belong to the class of enzymes called nucleases which degrade or cut single or double stranded DNA. A restriction endonuclease acts by recognizing and binding to particular sequences of nucleotides (the `recognition sequence') along the DNA molecule. Once bound, the endonuclease cleaves the molecule within or to one side of the recognition sequence. The location of cleavage may differ among various restriction endonucleases, though for any given endonuclease the position is fixed. Different restriction endonucleases have different affinity for recognition sequences. More than two hundred restriction endonucleases recognizing unique specificities have been identified among thousands of bacterial and archaeal species that have been examined to date.

Restriction endonucleases are classified on the basis of their composition and cofactor requirements, the nature of target sequence, and the position of the site of DNA cleavage with respect to the target sequence (Yuan, R. *Ann. Rev. Biochem.*, 50:285-315 (1981)). Currently three distinct, well-characterized classes of restriction endonucleases are known (I, II and III). Type I enzymes recognize specific sequences,

but cleave randomly with respect to that sequence. The type III restriction endonucleases recognize specific sequences, cleave at a defined position to one side of that sequence, but never give complete digestion. Neither of these two kinds of enzymes is suitable for practical use. The type II restriction endonucleases recognize specific sequences (4-8 nucleotides long) and cleave at a defined position either within or very close to that sequence. Usually they require only Mg^{2+} ions for their action. When they are purified away from other bacterial components, type II restriction endonucleases can be used in the laboratory to cleave DNA molecules into specific fragments. This property allows the researcher to manipulate the DNA molecule and analyze the resulting constructions.

Bacteria tend to possess at most, only a small number of restriction endonucleases per isolate. The restriction endonucleases are designated by a three-letter acronym derived from the name of organism in which they occur (Smith and Nathans, *J. Mol. Biol.* 81:419-423 (1973)). The first letter comes from the genus, and the second and third letters come from the species. Thus, a strain of the *species Deinococcus radiophilus* for example, synthesizes three different type II restriction endonucleases, named *Dra*I, *Dra*II and *Dra*III. These enzymes recognize and cleave the sequences TTTAAA, PuGGNCCPy and CACNNNGTG, respectively. *Escherichia coli* RY13, on the other hand, synthesizes only one type II restriction enzyme, *Eco*RI, which recognizes the sequence

GAATTC (Roberts R.J and Macelis D., *Nucl. Acids Res.*, 28:306-7 (2000)).

A second component of bacterial and archaeal restriction systems are the modification methylases (Roberts and Halford, in `Nucleases`, 2nd ed., Linn et al., ed.'s, p. 35-88 (1993)). These enzymes are complementary to restriction endonucleases and they provide the means by which bacteria are able to protect their own DNA and distinguish it from foreign, invading DNA. Modification methylases recognize and bind to the same recognition sequence as the corresponding restriction endonuclease, but instead of cleaving the DNA, they chemically modify one or other of the nucleotides within the sequence by the addition of a methyl group. Following methylation, the recognition sequence is no longer cleaved by the restriction endonuclease. The DNA of a bacterial cell is modified by virtue of the activity of its modification methylase, and is therefore insensitive to the presence of the endogenous restriction endonuclease. It is only unmodified, and therefore identifiably foreign DNA, that is sensitive to restriction endonuclease recognition and cleavage.

It is thought that in nature, type II restriction endonucleases cleave foreign DNA such as viral and plasmid DNA when this DNA has not been modified by the appropriate modification enzyme (Wilson and Murray, *Annu. Rev. Genet.* 25:585-627 (1991)). In this way, cells are protected from invasion by foreign DNA. Thus, it has been widely believed

that evolution of type II restriction modification systems has been driven by the cell's need to protect itself from infection by foreign DNA (the cellular defense hypothesis).

With the advent of genetic engineering technology, it is now possible to clone genes and to produce the proteins and enzymes that they encode in greater quantities than are obtainable by conventional purification techniques. The key to isolating clones of restriction endonuclease genes is to develop a simple and reliable method to identify such clones within gene libraries. One potential difficulty is that some restriction endonuclease and methylase genes may not express in *E. coli* due to differences in the transcriptional and translational machinery of the source organism and of *E. coli*, such as differences in promotor or ribosome binding sites or the codon composition of the gene. The isolation of the methylase gene requires that the methylase express well enough in *E. coli* to fully protect at least some of the plasmids carrying the gene. The isolation of the endonuclease in active form requires that the methylase express well enough to protect the host DNA fully, or at least enough to prevent lethal damage from cleavage by the endonuclease. Another obstacle to cloning restriction-modification systems lies in the discovery that some strains of *E. coli* react adversely to cytosine or adenine modification; they possess systems that destroy DNA containing methylated cytosine (Raleigh and Wilson, *Proc. Natl. Acad. Sci., USA* 83:9070-9074, (1986)), or methylated adenine (Heitman and Model, *J. Bact.* 196:3243-

3250, (1987)); Raleigh, et al., *Genetics*, 122:279-296, (1989)) Waite-Rees, et al., *J. Bacteriology*, 173:5207-5219 (1991)). Cytosine-specific or adenine-specific methylase genes cannot be cloned easily into these strains, either on their own, or together with their corresponding endonuclease genes. To avoid this problem it is necessary to use mutant strains of *E. coli* (McrA$^-$ and McrB$^-$ or Mrr$^-$) in which these systems are defective.

Several approaches have been used to clone restriction genes into *E. coli*:

1) Selection based on phage restriction

The first cloned systems used bacteriophage infection as a means of identifying or selecting restriction endonuclease clones (*Eco*RII: Kosykh et al., *Molec. Gen. Genet* 178:717-719, (1980)); *Hha*II: Mann et al., *Gene* 3:97-112, (1978)); *Pst*I: Walder et al., *Proc. Nat Acad. Sci.* 78:1503-1507, (1981)). Since the presence of restriction-modification systems in bacteria enable them to resist infection by bacteriophages, cells that carry cloned restriction-modification genes can, in principle, be selectively isolated as survivors from libraries that have been exposed to phage. This method has been found, however, to have only limited value. Specifically, it has been found that cloned restriction-modification genes do not always manifest sufficient phage

resistance to confer selective survival under standard conditions.

2) Selection based on vector modification

A second approach which is being used to clone a growing number of systems, involves selection for an active methylase gene (refer to U.S. Pat. No. 5,200,333 and *Bsu*RI: Kiss et al., *Nucl. Acid. Res.* 13:6403-6421, (1985)). Since restriction and modification genes are often closely linked, both genes can often be cloned simultaneously. This selection does not always yield a complete restriction system however, but instead may yield only the methyltransferase gene (*Bsp*RI: Szomolanyi et al., *Gene* 10:219-225, (1980); *Bcn*I: Janulaitis et al, *Gene* 20:197-204 (1982); *Bsu*RI: Kiss and Baldauf, *Gene* 21:111-119, (1983); and *Msp*I: Walder et al., *J. Biol. Chem.* 258:1235-1241, (1983)).

3) Sub-cloning of natural plasmids

Another cloning approach involves transferring systems initially characterized as plasmid-borne into *E. coli* cloning plasmids (*EcoRVI*: Bougueleret et al., *Nucl. Acid. Res.* 12: 3659-3676, (1984); *Pae*R7I: Gingeras and Brooks, *Proc. Natl. Acad. Sci. USA* 80:402-406, (1983); Theriault and Roy, *Gene* 19:355-359 (1982); *Pvu*II: Blumenthal et al., *J. Bacteriol.* 164:501-509, (1985)).

4) Multi-step cloning

Sometimes the straight-forward methylase selection method fails to yield a methylase (and/or endonuclease) clone due to various obstacles. See, e.g., Lunnen, et al., *Gene,* 74(1):25-32 (1988). One potential obstacle to cloning restriction-modification genes lies in trying to introduce the endonuclease gene into a host not already protected by modification. If the methylase gene and endonuclease gene are introduced together as a single clone, the methylase must protectively modify the host DNA before the endonuclease has the opportunity to cleave it. On occasion, therefore, it might only be possible to clone the genes sequentially, methylase first then endonuclease (see, U.S. Pat. No. 5,320,957).

5) Selection based on induction of the DNA-damage-inducible SOS response

Another method for cloning methylase and endonuclease genes is based on a colorimetric assay for DNA damage (see, U.S. Pat. No. 5,492,823). When screening for a methylase, the plasmid library is transformed into a sensitive *host E. coli* strain such as AP1-200. The expression of a methylase will induce the SOS response in an *E. coli* strain which is $McrA^+$, $McrBC^+$, or Mrr^+. The AP1-200 strain is temperature sensitive for the Mcr and Mrr systems and includes a *lacZ* gene fused to the damage inducible *dinD* locus

of *E. coli*. The detection of recombinant plasmids encoding a methylase or endonuclease gene is based on induction at the restrictive temperature of the *lacZ* gene. Transformants encoding methylase genes are detected on LB agar plates containing X-gal as blue colonies. (Piekarowicz, et.al., *Nucleic Acids Res.* 19:1831-1835, (1991) and Piekarowicz, et.al. *J. Bacteriology* 173:150-155 (1991)). Likewise, the *E. coli* strain ER1992 contains a dinD1-Lac Z fusion but is lacking the methylation dependent restriction systems McrA, McrBC and Mrr. In this system (called the "endo-blue" method), the endonuclease gene can be detected in the absence of it's cognate methylase when the endonuclease damages the host cell DNA, inducing the SOS response. The SOS-induced cells form deep blue colonies on LB agar plates supplemented with X-gal. (Fomenkov, et.al. *Nucleic Acids Res.* 22:2399-2403 (1994) and U.S. Pat. No. 5,498,535).

6) N-terminal-sequence-based degenerate inverse PCR method

It may occur that a modification methyltransferase gene cannot be identified (see, U.S. Pat. No. 5,945,288), or that a methylase gene can be identified but the open reading frame specifying the restriction endonuclease is uncertain. In these cases, an additional procedure for identifying the gene for the endonuclease specifically can be applied when the restriction endonuclease can be purified in sufficient quantity and purity from the original organism. In this method, the restriction

endonuclease is purified to substantial homogeneity and subjected to polypeptide sequencing. The polypeptide sequence obtained is reverse-translated into DNA sequence and degenerate PCR primers can be designed to amplify a portion of the endonuclease gene from genomic DNA of the original organism or from a gene library made therefrom. The DNA sequence of the complete genes can be obtained by methods dependent on Southern blot analysis or by further direct or inverse PCR methods. If the cognate methyltransferase gene cannot be obtained or cannot be expressed, the stability and utility of the solo restriction endonuclease clone will usually be severely compromised.

It may occur that genes for both the methyltransferase and the restriction endonuclease of a particular system can be obtained by the methods described above, but nevertheless establishment of a usable strain for enzyme production is problematic. Frequently the difficulty is with expression of the methyltransferase gene at a suitable level. This is particularly true with method (6). Such clones sometimes can be stabilized by using heterospecific methyltransferase genes, which were not associated with the endonuclease gene in the original host but which recognize the same or a related sequence and prevent the endonuclease from cleaving its recognition sequence (see, U.S. Pat. No. 6,048,731).

It may occur that there is no suitable heterospecific methyltransferase available, and the degree of protection conferred on the host by the cognate methyltransferase is inadequate; or it may occur that apparently adequate levels of methyltransferase can be obtained but such level is toxic to the cell, resulting in strains that cannot be stored; or it may occur that protection is apparently adequate and the protected strain is viable, but the combination of the methyltransferase and the endonuclease genes gives a strain that does not express detectable endonuclease; or it may occur that protection is apparently adequate, but the combination of the methyltransferase and the endonuclease genes gives a strain that expresses detectable endonuclease, but is not sufficiently stable to make commercially useful levels of enzyme.

Many factors can be imagined that might alter the requisite level of enzyme needed for effective protection of the host cell from cleavage by a restriction endonuclease. Such factors include rapid growth, during which more DNA copies are present in the cell than are present during the stationary phase of growth; recovery from a resting state, during which time new synthesis of the modification methyltransferase may be required before new synthesis of the restriction endonuclease begins; starvation of various sorts, during which time levels of required DNA methyltransferase cofactors such as S-adenosylmethionine may be altered; and special physiological states, such as DNA

damage or other physiological insults. In addition, levels of methyltransferase can potentially be too high and become toxic, for example by binding to or methylating extraneous sites related to the cognate site and thus interfering with the reading of the DNA sequence by regulatory or DNA-condensing proteins. Thus, the absolute level of expression of the methyltransferase may need to fluctuate in response to conditions over the life of a culture, in order to be indefinitely perpetuated.

This need for a fine level of control is not unique to modification methyltransferases. Over the course of 50 years of study, many detailed regulatory schemes have been described for various sorts of functions, such as catabolic and anabolic gene sets that break down nutrients (lac, ara, gal) or synthesize essential compounds (trp, his), or response to stressors (the DNA damage response, the heat shock response). These regulatory effects are mediated by changes in promoter activity (by activators or repressors), in transcript stability (by retroregulatory elements), by alteration of translation levels (by attenuation or translational coupling), for example. Despite this high level of understanding, it is not straighforward to anticipate in advance how the demand for a function will change with physiological changes and how to achieve the desired level of a function.

Because purified restriction endonucleases, and to a lesser extent, modification methylases, are useful tools for

characterizing genes in the laboratory, there is a commercial incentive to obtain bacterial strains through recombinant DNA techniques that synthesize these enzymes in abundance. Such strains would be useful because they would simplify the task of purification as well as providing the means for production in commercially useful amounts.

SUMMARY OF THE INVENTION

In one embodiment, the present invention relates to a method for cloning and expressing a target restriction modification system, comprising first implementing a method for producing a balanced level of activity of a protective modification methyltransferase, such that expression compensates for changes in the physiological state of the cell and therefore confers full protection preferably during all growth phases from cleavage by the cognate restriction endonuclease; and then introducing the restriction endonuclease gene and providing for its expression.

The invention further relates to the method for producing a balanced level of activity of a protective modification methyltransferase comprising specifically testing for the extent of protection during critical growth phases which may be selected from stationary phase, the logarithmic phase of growth, recovery from storage or other growth phases, and then identifying a suitable expression vector by selecting for its function at those critical growth phases.

The above method is exemplified by the cloning and expression of the MseI restriction modification system, which is encoded on a DNA (deoxyribonucleic acid) fragment, which fragment codes for two related enzymes, namely an enzyme which recognizes the DNA sequence 5'-TTAA-3' and cleaves the phosphodiester bond between the T residues of this recognition sequence to produce a 2 base 5' extension (Morgan R.D., *Nucl. Acids Res.,* 16:3104 (1988)) (hereinafter referred to as the *Mse*I restriction endonuclease), and a second enzyme, known as M.*Mse*I, that recognizes the same DNA sequence, 5'-TTAA-3', but modifies this sequence by the addition of a methyl group to prevent cleavage by the *Mse*I endonuclease. In addition, the invention relates to two additional DNA fragments, each of which encodes an enzyme differing in sequence from M.*Mse*I that perform the same function as M.*Mse*I, namely modifying the sequence 5'-TTAA-3' by the addition of a methyl group thus preventing cleavage by the *Mse*I endonuclease. The present invention also relates to a process for preparing the DNA fragment, a vector containing the DNA fragment, a transformed host containing this DNA fragment, and an improved process for producing *Mse*I restriction endonuclease from such a transformed host.

*Mse*I restriction endonuclease produced according to the present invention is substantially pure and free of the contaminants commonly found in restriction endonuclease preparations made by conventional techniques.

The *Mse*I methylase gene, but not the *Mse*I endonuclease gene, was obtained generally in accordance with the technique referred to as methylase selection (U.S. Pat. No. 5,200,333, the disclosure of which is hereby incorporated by reference herein). However none of the clones obtained by methylase selection expressed detectable *Mse*I restriction endonuclease activity and none was fully protected from *Mse*I digestion after overnight incubation. A methylase clone was sequenced and the *Mse*I methylase gene was identified based on homology to other N6-adenine methylases. Although the methylase clone did not produce any detectable *Mse*I endonuclease activity, it was speculated that the endonuclease gene was likely located adjacent to the methylase gene. DNA contiguous to the *Mse*I methylase gene was therefore amplified from *Micrococcus* species by inverse PCR techniques and sequenced.

To locate and positively identify the *Mse*I endonuclease gene, the N-terminal amino acid sequence of highly purified *Mse*I restriction endonuclease protein obtained from *Micrococcus* species was determined. An open reading frame in which the deduced amino acid sequence matched the N-terminal amino acid sequence of the *Mse*I endonuclease was observed in the DNA sequence obtained by inverse PCR techniques and located 3' of the methylase gene. The *Mse*I methylase gene was amplified and cloned into a vector compatible with a standard high expression vector. The *Mse*I

endonuclease gene was then amplified, ligated to an expression vector such as the pET series of vectors, and introduced into a host which was pre-modified with the *Mse*I methylase carried on a separate compatible vector; however, no *Mse*I activity was found in the few such constructs obtained. From further results below, it appears that this failure of *Mse*I expression from inadequate expression of the methylase so that successful endonuclease expression became a lethal event. After obtaining a fully modifying vector in accordance with the present invention, the expression of the endonuclease was also carefully regulated by construction of a vector which suppressed expression of the endonuclease during cell growth prior to the induction of the endonuclease gene. A host carrying the endonuclease and methylase genes in these special constructs was then grown, induced and harvested and used to make the *Mse*I endonuclease.

The preferred method for cloning and expressing the *Mse*I restriction-modification system consists of obtaining methylase positive clones according to methylase selection method and determining the DNA sequence of these *Mse*I methylase positive clones. The DNA adjacent to the methylase gene is obtained by inverse PCR techniques and sequenced. The *Mse*I endonuclease protein from *Micrococcus species* is purified to near homogeneity and the N-terminal amino acid sequence determined. The *Mse*I endonuclease gene is identified based on the DNA sequence and amino acid

sequence data. The expression of the *Mse*I methylase is modulated to achieve full protection of the host genome without creating so much methylase expression as to be toxic to the host. This full methylation state is monitored by testing DNA obtained from cells in rapid logarithmic growth for protection from *Mse*I endonuclease cleavage and using a construct which provides full protection under these rapid growth conditions. The *Mse*I endonuclease is then expressed by amplifying the complete gene from *Micrococcus* species genomic DNA and ligating it into an expression vector designed to limit expression of the *Mse*I endonuclease during cell growth prior to induction, such as pVR-24 (New England Biolabs, Inc., Beverly, Mass.). The construct is introduced into a host with appropriate genetic composition to provide sufficient regulatory capacity (U.S. Application Serial No. _____) which is premodified at *Mse*I sites by virtue of carrying the *Mse*I methylase gene expressed on the separate compatible plasmid providing full protection against *Mse*I cleavage. The *Mse*I endonuclease is produced by growing the host containing the *Mse*I endonuclease and methylase genes, inducing with the appropriate expression conditions, harvesting the cells and purifying the *Mse*I endonuclease therefrom.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1A shows a restriction map of the recombinant plasmid pVR-18 encoding *Mse*I DNA methyltransferase gene.

Figure 1B shows the agarose gel analysis of the susceptibility to *Mse*I of pVR-18 plasmid encoding M. *Mse*I. Lane 1, uncut pVR-18; lane 2, pVR-18 following overnight incubation with ten units of *Mse*I; lane 3, uncut pBR322; lane 4, pBR322 following overnight incubation with ten units of *Mse*I; lane 4, pVR-18 + pBR322 following overnight incubation with ten units of *Mse*I; lane 5, molecular weight standard (1 kb DNA Ladder, New England Biolabs, Inc.).

Figure 2A shows a restriction map of the recombinant plasmid pEsaDix4I encoding putative DNA methyltransferase gene.

Figure 2B shows the agarose gel analysis of the susceptibility to *Mse*I of pEsaDix4I plasmid encoding a putative DNA methyltransferase gene. Lane 1, uncut pEsaDix4I; lane 2, pEsaDix4I following overnight incubation with ten units of *Mse*I; lane 3, uncut pUC19; lane 4, pUC19 following overnight incubation with ten units of *Mse*I; lane 4, pEsaDix4I + pUC19 following overnight incubation with ten units of *Mse*I; lane 5, molecular weight standard (1 kb DNA Ladder, New England Biolabs, Inc.).

Figure 3A shows a restriction map of the recombinant plasmid pEsaDix5I encoding putative DNA methyltransferase gene.

Figure 3B shows the agarose gel analysis of the susceptibility to *Mse*I of pEsaDix5I plasmid encoding putative DNA methyltransferase gene. Lane 1, uncut pEsaDix5I; lane 2, pEsaDix5I following overnight incubation with ten units of *Mse*I; lane 3, uncut pUC19; lane 4, pUC19 following overnight incubation with ten units of *Mse*I; lane 4, pEsaDix5I + pUC19 following overnight incubation with ten units of *Mse*I; lane 5, molecular weight standard (1 kb DNA Ladder, New England Biolabs, Inc.).

Figure 4 shows the DNA sequence of *mseIM* gene (SEQ ID NO: 1) and its encoded amino acid sequence (SEQ ID NO: 2).

Figure 5 shows the DNA sequence of *esaDix4IM* gene (SEQ ID NO:3) and its encoded amino acid sequence (SEQ ID NO:4).

Figure 6 shows the DNA sequence of *esaDix5IM* gene (SEQ ID NO:5) and its encoded amino acid sequence (SEQ ID NO:6).

Figure 7 shows the DNA sequence of *mseIR* gene (SEQ ID NO:7) and its encoded amino acid sequence (SEQ ID NO:8).

Figure 8 shows a restriction map of the recombinant plasmid pNKR1707*mseIM* used for construction of a library of

constitutive promoters randoming mutagenized by error-prone PCR.

Figure 9A shows a restriction map of the recombinant plasmid pNKR1707mseIM-9 encoding the *Mse*I DNA methyltransferase gene and upstream regulatory elements.

Figure 9B shows the DNA sequence upstream of *Mse*I DNA methyltransferase gene (SEQ ID NO:9) which contains an optimal promoter sequence.

Figure 10 shows the construction of the plasmids pVR-26 and pVR-27 used for controlled expression of *Mse*I DNA methyltransferase gene.

Figure 11 shows the construction of the pVR-24 expression vector.

Figure 12A shows a restriction map of pVR-25 encoding the *Mse*I restriction endonuclease gene.

Figure 12B shows the mechanism of action of the tight regulatory system in pVR-25 for cloning genes encoding cytotoxic proteins.

Figure 13 shows a restriction map of pCEF-8 encoding T7 lysozyme gene.

Figure 14 shows an assay of *Mse*I restriction endonuclease activity in crude cell extracts made from *E. coli* strains MseRM4, MseRM5 and MseRM6. The growth conditions are described in Example IV.

Figure 15 shows an assay of *Mse*I restriction endonuclease activity in crude cell extracts made from *E. coli* strain MseRM4 (NEB #1284)after growth in the 100-liter fermenter.

DETAILED DESCRIPTION OF THE INVENTION

In one embodiment, the present invention relates to a method of producing a target restriction endonuclease by first providing a vector expressing a modification methyltransferase gene protecting DNA from restriction enzyme cleavage, in such a form that complete protection of the host DNA is observed preferably at all growth phases in which the cognate restriction endonuclease is present without leading to toxicity (a fully-protecting methyltransferase vector), followed by providing a vector expressing the desired restriction endonuclease gene. The present invention is not limited by the identity of the modification methyltransferase gene or restriction enzyme, except that the modification methyltransferase must protect against cleavage by the said restriction enzyme.

In a preferred embodiment, the fully-protecting methyltransferase vector may be obtained by identifying regulatory elements capable of driving methyltransferase expression to provide full protection during a phase of growth that is especially sensitive to methyltransferase expression pattern. In accordance with the present invention, this may be done by the following steps:

(1) obtaining a methyltransferase gene in a vector by methods known in the art;

(2) placing a regulatory element such as a promoter in a suitable location with respect to the gene;

(3) transformation into the desired host cell;

(4) reisolation of vector from the pooled transformants during the time that they are in the logarithmic phase of growth;

(5) selection by digestion with the endonuclease; and

(6) retransformation of the surviving undigested and thus protected vector population into a fresh host.

It will be understood by those skilled in the art that step (4) of this procedure may be performed with pooled vector isolated from logarithmic phase or from various other phases of growth, for example from stationary phase, from a resting state achieved by starvation for carbon or nitrogen or other essential nutrient, or from cells in a special physiological state, such as a state of DNA damage, or in the presence of

physiological insults such as acidic media or toxic compounds, as may be appropriate.

In a preferred embodiment, the regulatory element of step (2) is identified by a procedure comprising the following steps:

(a) cloning into the vector containing the methyltransferase gene, at a desired location, a pool of fragments containing various distinct regulatory elements; and

(b) proceeding with steps (3) through (6).

It will further be understood by those skilled in the art that the process of selection comprising steps (3) through (6) may be repeated to select further improvement.

It will further be understood by those skilled in the art that step (2a), cloning a pool of fragments containing regulatory elements at the same or a different location, may be repeated followed by repeated selection as may be appropriate.

The present invention is not limited by the identity of the regulatory element, which may be a promoter, an operator, an enhancer, or a down-stream regulatory element.

In a preferred embodiment, the methyltransferase gene of step (1) is isolated by the methylase selection procedure

(U.S. Pat. No. 5,200,333). The present invention is not limited to methyltransferase genes isolated in this way but includes genes isolated by any of the methods described above such as phage selection, subcloning of natural plasmids, identification based on induction of the DNA-damage-inducible SOS response, by inverse PCR based on amino acid sequence of a purified protein, or identification in sequence databases from similarity to sequences of other methyltransferase followed by cloning by PCR or by Southern blot based procedures (see e.g., Kong, et al., *Nucleic Acids Res.* 28:3216-3223 (2000)).

In a preferred embodiment, the collection of distinct regulatory elements of step (2a) comprises copies of the *his* promoter of *S. typhimurium* randomly mutagenized by error-prone PCR together with such contaminating chromosomal fragments as may be present in the preparation of mutagenized fragments. The present invention is not limited to fragments obtained in this way, but may include collections of fragments isolated from genomic DNA of *E. coli* or another organism or fragments derived by oligonucleotide synthesis with degenerate sequences at random or specific locations or fragments derived by recombinational PCR of a random or specific collection of fragments. In a preferred embodiment, the regulatory element obtained in this way is the sequence of SEQ ID NO:9.

It will further be understood by one skilled in the art that this method may be applied to any methyltransferase that confers protection from cleavage by the restriction endonuclease in question, not merely that which co-occurs with the said endonuclease in a particular natural isolate.

The present invention further relates to the isolation of methyltransferase genes of desired specificity from DNA of environmental sources without first culturing the organisms contained therein. In a preferred embodiment, these genes are isolated by methylase selection from DNA made from a sample of a mixed green filament and mat community of prokaryotes growing at 68°C at Dixie Valley Hot Spring, Nevada.

The present invention further relates to provision of a desired restriction endonuclease gene expressed from a vector with tight regulation such that extremely low levels of protein are expressed in the absence of induction (very low basal expression is observed). In a preferred embodiment this tight regulation vector comprises a vector with antagonistic and independently regulatable promoters reading through the cloned target gene as described in WO 99/11821 and U.S. Application Serial No. 09/486,356 (now U.S. Patent 6,383,770) but basal expression has been further lowered by providing for a lower copy number than is present in the previously existing vector pLT7K used for this purpose. In a preferred embodiment, the copy number of the vector is lowered by

10/21/03

exchanging the replication origin of pLT7K for that of pACYC184. Other replication origins might also be used, such as those of pSC101 (Stoker, et al., *Gene* 18:335-341 (1982)), pSYX20 (U.S. Patent No. 5,262,313), F (Shizuya, et al., *Proc. Natl. Acad. Sci. USA* 89(18):8794-8797 (1992)) or other low-copy vectors (Harayama, et al., *Mol. Gen. Genet*. 184:52-55 (1981) and Wohlfarth, et al., *J. Gen. Microbiol.* 134:433-440 (1988)). In a preferred embodiment, the vector is pVR-24.

In a preferred embodiment, further lowering of basal expression level is achieved by employment of a strain expressing high levels of the negative regulator of expression in the direction that allows translation of the target gene, as described in the accompanying U.S. Application Serial No. _____.

The above described method is exemplified in another embodiment of the present invention, namely the cloning and expression of the *Mse*I restriction-modification system.

The present invention also provides novel DNA constructs and novel compositions comprising microbial strains producing *Mse*I restriction endonuclease. The restriction endonuclease of interest in the present invention, *Mse*I, recognizes the DNA sequence 5'-TTAA-3' and cleaves the phosphodiester bond on between the T residues of this recognition sequence to produce a 2 base 5' extension.

In order to overexpress the *Mse*I restriction endonuclease, additional steps beyond the well-known art of the methylase selection procedure (U.S. Pat. No. 5,200,333) are required, including particularly the fine balance of *Mse*I methyltransferase expression to fully protect the host genomic DNA from *Mse*I digestion *in vivo* while yet not producing so much methyltransferase as to be toxic to the host. A vector, containing the *mseIM* gene optimized for expression such that full protection against *Mse*I endonuclease is observed even during very rapid (logarithmic stage of) cell growth, is first used to modify an *E. coli* host. This host is then transformed with a compatible vector, such as pVR-25, containing the *mseIR* gene followed by selection for colonies that contain both vectors on the appropriate antibiotic plates. *Mse*I endonuclease producing constructs are identified by growing individual transformants and assaying for *Mse*I endonuclease activity, (as in Example V below).

The method described herein by which the *Mse*I methylase gene and the *Mse*I restriction endonuclease genes are preferably cloned and expressed in *E. coli* employs the following steps:

1) Cloning of the DNA methyltransferase genes which protect from *Mse*I cleavage.

It is well known that DNA modification methylases recognize and bind to the same nucleotide recognition

sequence as the corresponding restriction endonuclease, but instead of breaking the DNA, they chemically modify one or other of the nucleotides within the sequence by the addition of a methyl group. Following this methylation, the recognition sequence is no longer bound or cleaved by the restriction endonuclease. The DNA of a bacterial cell is always fully modified, by virtue of its modification methylase, and it is therefore completely insensitive to the presence of the endogenous restriction endonuclease. In this situation, only unmodified, and therefore identifiably foreign, DNA that is sensitive to restriction endonuclease recognition and attack. The first step of present method is to identify the DNA methyltransferase gene which protects from *Mse*I cleavage. To accomplish this the DNA methylase from *Micrococcus species* (NEB446) can be cloned. Alternatively, a DNA methyltransferase from an R-M system other than the *Mse*I R-M system, but able to protectively modify DNA to prevent digestion by the *Mse*I restriction enzyme can be identified as described in U.S. Pat. No. 5,179,015. In the present invention, three DNA methylases able to protect DNA from digestion by the *Mse*I restriction enzyme were identified.

First, the total genomic DNA was purified from *Micrococcus species* (NEB#446). A random library of this DNA was constructed by partially digesting the DNA with a frequent cutting endonuclease, *Sau*3AI, to produce fragments of approximately 1 to 10 kilobases (kb) in length. These fragments were ligated into a vector pBR322, previously

cleaved with *Bam*HI and dephosphorylated. The ligation reaction was transformed into chemically competent *E. coli* ER2502 cells. The transformants were pooled, and the plasmid was population purified to form the primary plasmid library. An aliquot of these purified plasmids was digested with *Mse*I restriction endonuclease to destroy all plasmids which had not expressed the *Mse*I methylase gene *in vivo* and thus protected the plasmid DNA from digestion. The digested plasmid pool was transformed again into *E. coli* ER2502 to recover the intact, *Mse*I methylase expressing plasmids. Individual clones were picked, there plasmid DNA was purified and challenged by cleavage with *Mse*I endonuclease. Plasmids which were not cut by *Mse*I contained the *Mse*I methyl-transferase gene.

In a preferred embodiment, the methyltransferase gene is one protecting against the restriction endonuclease *MseI* obtainable from *Micrococcus species* (NEB#446), and may be selected from among the set of sequences that can encode proteins specified in SEQ ID NO:2, SEQ ID NO:4, and SEQ ID NO:7. These proteins may be encoded for example by those DNA sequences set forth in SEQ ID NO:1, SEQ ID NO:3, and SEQ ID NO:6.

To search for alternative DNA methyltransferases that are able to protect DNA from cleavage by *Mse*I endonuclease, a library of clones from a source of DNA other than *Micrococcus* species may be constructed in a vector containing one or

more *Mse*I restriction sites. This library of clones is then selected by one or more rounds of *Mse*I digestion to destroy non-protecting clones followed by transformation of the digested plasmids to recover protected clones. Such a library was created from DNA (designated "environmental DNA") isolated from a sample of a mixed green filament and mat community of prokaryotes growing at 68°C at Dixie Valley Hot Spring, Nevada. Purified enviromental DNA was digested with *Nsi*I endonuclease and ligated into the vector pNEB193 previously cleaved with *Pst*I restriction endonuclease and dephosphorylated. The ligation reaction was transformed into *E. coli* ER2683 by electroporation. The transformants were pooled and the plasmid population was purified to form the primary plasmid library. An aliquot of these purified plasmids was digested to completion with an excess of *Mse*I restriction endonuclease and used to transform ER2683. Plasmids of the resulting transformants were miniprepped and analyzed by *Mse*I restriction enzyme digestion and subsequent agarose gel electrophoresis. 9 plasmids examined were found to be resistant to *Mse*I digestion and each was found to encode one of two different methylase genes that each function to protect DNA from cleavage by *Mse*I. These two methylases were named *esaDix4IM* (SEQ ID NO:3 and SEQ ID NO:4) and *esaDix5IM* (SEQ ID NO:5 and SEQ ID NO:6). Analysis of crude cell extracts prepared from these clones revealed no endonuclease activity. These methyltransferases, or others like them, may be used to protect a host's own DNA and thus enable the successful expression of the *Mse*I endonuclease.

2) Sequence determination of the entire *Mse*I restriction-modification system.

The *Mse*I methylase gene, but not the *Mse*I endonuclease gene, was obtained generally in accordance with the technique referred to as methylase selection (U.S. Pat. No. 5,200,333) as above in step 1. However none of the clones obtained by methylase selection expressed detectable *Mse*I restriction endonuclease activity. A methylase clone was sequenced using standard techniques on an ABI 373 DNA sequencing machine. The *Mse*I methylase gene was identified based on amino acid homology to other N6-adenine methylases. Although the methylase clone did not produce any detectable *Mse*I endonuclease activity, it was speculated that the endonuclease gene was likely located adjacent to the methylase gene. DNA contiguous to the *Mse*I methylase gene obtained from *Micrococcus* species (NEB#446) was therefore amplified from *Micrococcus* species genomic DNA by inverse PCR techniques and sequenced.

To locate and positively identify the *Mse*I endonuclease gene, the N-terminal amino acid sequence of highly purified *Mse*I restriction endonuclease protein obtained from *Micrococcus* species was determined. *Mse*I endonuclease may be purified from *Micrococcus* species (NEB#446) as set forth in Example III below. An open reading frame in which the deduced amino acid sequence matched the N-terminal amino

acid sequence of the *Mse*I endonuclease was observed in the DNA sequence obtained by inverse PCR techniques which was located 3' of the methylase gene.

Alternatively, the N-terminal amino acid sequence of *Mse*I restriction endonuclease can be used to design degenerate oligonucleotide primers for PCR amplification of a portion of the *Mse*I endonuclease gene from *Micrococcus* species (NEB#446). The DNA sequence obtained can then be used to guide inverse PCR amplification of the DNA on either side of this original portion of the *Mse*I endonuclease gene, and the *mseIM* and *Mse*I genes can be identified in this DNA sequence as above. Both methods were used for cloning and sequence determination of the entire *Mse*I restriction-modification system.

3) Fine optimization of the *Mse*I methyltransferase expression

Once the complete genes for the *Mse*I endonuclease and *Mse*I methyl-transferase have been identified (SEQ ID NO:7, SEQ ID NO:8 and SEQ. ID. NO:1 and SEQ ID NO:2, respectively), they may then be manipulated in a variety of ways to provide for expression. Using the methylase constructs obtained as above, expression of the *Mse*I restriction endonuclease gene under T7 promoter control using the pET series of vectors (Novagen Inc., Madison, WI)

was vigorously attempted but failed to yield a *Mse*I restriction endonuclease producing clone.

A unique combination of methods, including the introduction of a second, controllable promoter before the methylase gene, using a low copy replicon for the endonuclease gene and increasing the copy number of LacI repressor in the host prior to the introduction of the endonuclease gene, was used to control the overexpression of recombinant *Mse*I endonuclease.

It was observed that the methylase constructs obtained by methylase selection did not fully protect the host *E. coli* chromosomal DNA when the cells were rapidly growing in logarithmic phase of growth. In order to increase expression of the methylase, and thus fully protect the host DNA so that *mseIR* could be introduced successfully into the cells and expressed, the methylase gene was amplified from *Micrococcus species* DNA and cloned into a family of vectors (pNK series, see Example IV below) under the expression of various strength constitutive promoters. In this attempt, no methylase constructs were obtained for the two highest level of expression promoters, due we believe to toxicity to the cell from too much expression of the methylase. Constructs with the two lower level of expression promoters failed to fully protect the host against *Mse*I cleavage when checked at logarithmic phase of growth. In order to increase methylase expression to fully protect the host DNA during rapid growth

but remain below the level of toxicity, one of the promoter constructs was subjected to random mutagenesis by error-prone PCR in the promoter region. Mutated clones expressing *Mse*I methylase were selected using the methylase selection technique referenced above, and then individual clones were tested for the ability to fully protect host genomic DNA from *Mse*I cleavage during rapid logarithmic growth by harvesting cells during logarithmic growth, purifying DNA from these host cells and testing for full protection from *Mse*I cleavage. One of the constructs found to fully protect against *Mse*I was then used for the expression of the *Mse*I endonuclease.

This method of modulating expression of a methyltransferase to achieve full protection during all stages of host cell growth may prove applicable to other systems where the endonuclease proves difficult to express, or express instability in a host cell (see, U.S. Patent No. 6,025,179 and 6,048,731).

4) Expression of the *Mse*I restriction endonuclease under the control of an inducible promoter

To optimize expression of recombinant *Mse*I of the present invention, inducible or constitutive promoters are well known and may be used to express high levels of an *mseIR* gene in a recombinant host. Similarly, high copy number vectors, well known in the art, may be used to achieve high levels of expression. In accordance with the present

invention, it has been found that a particularly preferred method for expression of *Mse*I restriction endonuclease is an expression vector designed to limit expression of the *Mse*I endonuclease during cell growth prior to induction, such as pVR-24 (New England Biolabs, Inc., Beverly, Mass.). This plasmid contains the segment encoding replicative function (ori), a chloramphenicol-resistance gene (Cm), gene encoding kanamycin resistance which is flanked by restriction endonuclease sites suitable for cloning. The cI857 gene encodes a mutant form of the of the lambda bacteriophage repressor protein, which conditionally binds to DNA sequences (the CI operator) that overlap PL and PR (the lambda bacteriophage major leftward and rightward promoters, respectively). The *lacI* gene encodes a repressor protein, LacI, that conditionally binds a DNA sequence (the lac operator) which has been constructed to overlap P_{T7} (bacteriophage T7 RNA polymerase transcriptional promoter). Briefly, at high temperature (42°C) without IPTG, the antisense promoter is active, while P_{T7} is repressed by LacI. At 30°C and with IPTG expression occurs from P_{T7} (see Figures 11 and 12). At intermediate temperatures and with intermediate IPTG concentrations, intermediate levels of expression can be obtained.

To obtain a stable clone which overexpresses the restriction endonuclease, the host is generally pre-protected from restriction endonuclease digestion. In the present invention this is accomplished by cloning the *Mse*I methylase

gene, or another methylase gene that protects against *Mse*I cleavage, such as *esaDix4IM* or *esaDix5IM*, expressed on the separate compatible plasmid in a manner providing full protection against *Mse*I cleavage. As shown in the Example V below it was found that the stability of the expression plasmid containing the restriction endonuclease gene construct and/or its mRNA could be improved when the *Mse*I methyltransferase gene is preceded by a DNA fragment encoding a novel promoter sequence. The *Mse*I endonuclease is produced by growing the host containing the *Mse*I endonuclease and the protective methylase gene, inducing with the appropriate expression conditions, harvesting the cells and purifying the *Mse*I endonuclease therefrom.

The invention further provides a process for producing the *Mse*I restriction endonuclease, in which recombinant DNA modification methods are used for transforming a microorganism such that the gene encoding the *Mse*I restriction endonuclease and a gene coding for a DNA methyltransferase which protects the host DNA from *Mse*I cleavage are introduced into said microorganism, the organism is grown under conditions suitable for expression of *Mse*I endonuclease, harvested and the *Mse*I endonuclease is purified therefrom.

Although the above-outlined steps represent the preferred mode for practicing the present invention, it will be apparent to those skilled in the art that the above described

approach can vary in accordance with techniques known in the art.

The following Examples are given to illustrate embodiments of the present invention as it is presently preferred to practice. It will be understood that these Examples are illustrative, and that the invention is not to be considered as restricted thereto except as indicated in the appended claims.

The references cited above and below are herein incorporated by reference.

EXAMPLE I

Cloning of the *MseI* methyltransferase gene (*mseIM*).

Micrococcus species (NEB#446) was grown overnight in 1L of LB broth, the cells were harvested and genomic DNA was isolated using Qiagen Genomic-tip 100/G Genomic DNA Purification Kit (Cat. No. 10243) according to the manufacturer's instructions. Genomic DNA was partialy digested with *Sau*3AI to produce fragments from 1 to 10 kb, and 20 ug of this cleaved genomic DNA was ligated with 3 ug of *Bam*HI-digested and dephosphorylated pBR322. The ligation mixture was transformed into *E. coli* strain ER2502. Approximately 100,000 transformants were obtained. The transformants were pooled, grown in 500 ml LB broth containing 100μg/ml ampicillin, and the plasmid population

was purified to form the primary plasmid library. 2 micrograms of this plasmid library was digested to completion with an excess of *Mse*I restriction endonuclease and used to transform ER2505. Plasmids of the resulting transformants were subjected to a second round of selection. 80 transformants were obtained and the plasmid DNA of 16 of these was analyzed by *Mse*I restriction enzyme digestion and subsequent agarose gel electrophoresis. 14 out of 16 plasmids examined were found to be resistant to *Mse*I digestion and found to carry the same *mseIM* gene (SEQ ID NO:1, SEQ ID NO:2) on a *Sau*3AI fragment of approximately 1.6 kb. Analysis of crude cell extracts prepared from those 14 clones revealed no *Mse*I activity.

EXAMPLE II

Cloning two DNA methylases from an environmental DNA sample that protect DNA from cleavage by *Mse*I endonuclease .

To search for alternative DNA methyltransferases that are able to protect DNA from cleavage by *Mse*I endonuclease, a library of clones from a source of DNA other than *Micrococcus* species (NEB446) may be constructed in a vector containing one or more *Mse*I restriction sites. This library of clones is then selected as above by one or more rounds of *Mse*I digestion to destroy non-protecting clones followed by transformation of the digested plasmids to recover protected clones, as in Example I above. Such a library was created from DNA isolated from a sample of a mixed green filament

and mat community of prokaryotes growing at 68°C at Dixie Valley Hot Spring, Nevada. 2 micrograms of the DNA was digested with *Nsi*I endonuclease and ligated into 1 microgram of the vector pNEB193 previously cleaved with *Pst*I and dephosphorylated. The ligation reaction was transformed into *E. coli* ER2683 by electroporation and approximately 1,000,000 transformants were obtained. The transformants were pooled, grown in 500 ml LB broth containing 100μg/ml ampicillin, and the plasmid population was purified to form the primary plasmid library. 1 microgram of this plasmid library was digested to completion with an excess of *Mse*I restriction endonuclease and used to transform ER2683. Plasmids of the resulting transformants were miniprepped and analyzed by *Mse*I restriction enzyme digestion and subsequent agarose gel electrophoresis. 9 plasmids examined were found to be resistant to *Mse*I digestion and were found to encode one of either two different methylase genes that both function to protect DNA from cleavage by *Mse*I. These two methylases were named esaDix4IM and esaDix5IM (SEQ ID NO:3, SEQ ID NO:4 and SEQ ID NO:5 and SEQ ID NO:6). Analysis of crude cell extracts prepared from these clones revealed no endonuclease activity. These methyltransferases, or others like them, may be used to protect a host's own DNA and thus enable the successful expression of the *Mse*I endonuclease.

EXAMPLE III

Identification and sequence determination of the *Mse*I restriction endonuclease gene using N-terminal amino acid sequence and DNA sequence adjacent to the *Mse*I methylase obtained by the inverse PCR method.

A) Purification of the *Mse*I restriction endonuclease from *Micrococcus species* to near homogeneity:

Micrococcus species (NEB#446) cells were propagated in LB media at 30°C. The cells were harvested by centrifugation after 20 hours growth and stored at -70°C until used. All of the procedures were performed on ice or at 4°C. The *Mse*I endonuclease was purified following the same scheme as in Example VI. Approximately 10,000 units of *Mse*I activity were purified to near homogeneity. 16 µl of the peak fraction was loaded onto an SDS-PAGE protein gel and subjected to electrophoresis. The gel was stained with Coomassie blue R-250 and a prominent band at approximately 21 kD corresponding to the *Mse*I restriction endonuclease activity was observed.

B) Amino Terminal *Mse*I protein sequence:

The *Mse*I restriction endonuclease, prepared as described, was subjected to electrophoresis and electroblotted according to the procedure of Matsudaira (Matsudaira, P., *J. Biol. Chem*. 262:10035-10038 (1987), with modifications as previously described (Looney, et al., *Gene* 80:193-208 (1989)). The membrane was stained with

Coomassie Blue R-250 and the protein band of approximately 21 kd was excised and subjected to sequential degradation on an Applied BioSystems Division, Perkin-Elmer Corporation (Foster City, California) Model 407A gas phase protein sequencer (Waite-Rees, et al., *J. Bacteriol.* 173:5207-5219 (1991)). The first 25 residues of the 21 kD protein corresponded to (Met)-Thr-His-Glu-Pro-Thr-Asp-Asp-Pro-Asp-Phe-Ile-Val-Met-Ala-Ala-Ser-Ala-Xxx-Asn-Leu-Ala-Asp-Xxx-Tyr (SEQ ID NO:10). This data was used to compare with amino acid sequence deduced from the DNA sequence adjacent to the methylase gene to identify the endonuclease gene.

C) DNA sequence determination adjacent to the *mseIM* methylase:

Template preparation for inverse PCR amplification: 1 μg of *Micrococcus species* (NEB#446) DNA was digested with 10 units of *Hae*II restriction endonuclease in 1X NEBuffer #4 in a 50 μl reaction volume for 1 hour at 37°C. The *Hae*II enzyme was heat inactivated by incubating at 75°C for 20 minutes. The *Hae*II digested DNA was circularized by adding 50 μl 10X T4 DNA ligase buffer and 400 μl dH_2O, followed by 5 μl (2000 NEB units) T4 DNA ligase (NEB#202) and incubating at 16°C for 16 hours. A portion of this circularization ligation reaction was then used as the template for subsequent inverse PCR reactions.

Primers MseI-IP1 and MseI-IP2 of sequences shown below were synthesized. These primers hybridize within the *Mse*I endonuclease gene and are oriented in the opposite direction relative to each other.

Primer MseI-IP1

5'-CTTCTGCAGCCGATTTCATAGTGATGGC -3' (SEQ ID NO:11)

Primer MseI-IP2

5'- GTTCTGCAGATCGGGATCATCCGTCGG -3' (SEQ ID NO:12)

In the reaction that was successful in amplifying the product, a reaction mix was made by combining:

10 µl of 10X Vent® reaction buffer

6 µl of 4 mM dNTP solution

5 µl of primer MseI-IP1 at 10 µM concentration

5 µl of primer MseI-IP2 at 10 µM concentration

3 µl of 100 mM $MgSO_4$ (5 mM Mg^{++} final concentration)

12.5 µl of circularized DNA template (aproximately 25 ng)

58 µl dH_2O

2 µl (4 units) of Vent® Exo^- polymerase NEB#257

The PCR amplification conditions were: 95°C for 3 minutes for one cycle, followed by 4 cycles of 95°C for 30 seconds, 52°C for 30 seconds and 72°C for 1.5 minutes, followed by 20 cycles of 95°C for 30 seconds, 62°C for 30

seconds and 72°C for 1.5 minutes. 10 µl of the PCR reaction was analyzed by electrophoresis on a 0.8 % agarose gel.

An approximately 1350 bp product was observed in the *Hae*II circular template PCR reaction. The product was gel purified and suspended in 25 µl DNA (1X TE) buffer. This PCR product was then sequenced on an ABI 373 automated sequencing system according to the manufacturer's instructions, using the PCR primers above as the sequencing primers. Additionally, the *Mse*I endonuclease region was PCR amplified in a like reaction with the following primers and the PCR product was sequenced.

Primer MseI-IP3
5'-GGTTCTGCAGTTAAGGAGGTTTAACATATGACCCACGAACCGACG GATG-3' (SEQ ID NO:13)

Primer MseI-IP4
5'-GTTGGATCCGTCGACGCTTCTCGGCGTACCGAGCG-3'
(SEQ ID NO:14)

The *Mse*I endonuclease gene is identified by comparing the amino acid translation of DNA sequences adjacent to the *Mse*I methylase gene with the amino acid sequence data obtained from N-terminal amino acid sequencing of the *Mse*I endonuclease. An open reading frame oriented in the same direction as the *Mse*I methylase gene and overlaping the methylase gene by 7 amino acid residues was found in which

the first 25 amino acids coded for in the DNA sequence matched the amino acid sequence determined from the *Mse*I endonuclease protein.

Sequencing of the insert carrying the *Mse*I methylase and restriction endonuclease genes was performed using GPS®-1 of the Genome Priming System (New England Biolabs, Beverly, Mass.). GPS®-1 contains a modified Tn7 with the nptII gene for resistance to kanamycin, and insertions were generated in vitro in pVR-18 and pNEB193 containing part of the *Mse*I methylase gene according to the instructions of the manufacturer (New England Biolabs, Beverly, MA). These insertions were then sequenced using an ABI 373 automated sequencing system according to the manufacturer's instructions, using the primers included in GPS®-1 kit (Primer S and Primer N for the left and right end of the Transprimer, respectively)

EXAMPLE IV

Optimization of the *Mse*I M expression

1) Placing the *Mse*I methylase gene under different strength of constitutive promoters

To achieve a range of constitutive expression of the *Mse*I methylase, a related family of pNK vectors (generous gifts from N. Kleckner) containing constitutive promoters of different strengths was utilized. These plasmids contain either

the WT or mutated pHis promoters upstream of a *Bam*HI site and are derivatives of the RS415 plasmid (Simons, et al. *Gene*, 53 (1987) 85-96). Their designations and promoter strength are as follows:

No.	Plasmids	Promoter Strength
1	pNK1707 (wildtype)	1x
2	pNK2213	20x
3	pNK1786	100x
4	pNK2138	1070x

The above plasmids were digested by *Bam*HI, *Mun*I and *Ban*II and the vector backbones containing the constitutive promoters were gel purified. (The *Ban*II digest was included to aid in gel purification of the vector backbone by eliminating a similarly sized plasmid fragment.)

To prepare the *Mse*I methylase gene for insertion downstream of the constitutive promoters described above, PCR was done using Vent® DNA polymerase, 1X ThermoPol buffer, 4 mM $MgSO_4$, 80 ng of pVR19 plasmid (R. Vaisvila) containing the *Mse*I methylase gene as the template in a 100 µl PCR reaction, and primers introducing an upstream *Bam*HI site 5'-GAACCGGATCCGACCCTGAGTGAGAACATGCC-3' (SEQ ID NO:15) and a downstream *Mfe*I site 5'-AGGTCGCAATTGCCAGG GGTCGTCTTCACTCGCTAC-3' (SEQ ID NO:16) with respect to the methylase gene. Twenty-five cycles were done consisting of 10 sec at 95°C, 60 sec at 60°C and 75 sec at 72°C. The resulting 1019 bp PCR product was purified using a QiaQuick

PCR purification protocol, digested sequentially by *Bam*HI and *Mun*I, and purified once again using the QiaQuick PCR purification protocol.

The *Mse*I methylase gene was ligated into all four *Bam*HI-*Mun*I vector backbones, transformed into ER2688 cells, and plated on Luria-Bertani (supplemented with 1 gram glucose and 1 gram $MgCL_2$ per liter; subsequently referred to as supplemented LB) agar plates. However, attempts to place the *Mse*I methylase under the highest two levels of expression failed, assumingly due to instability from high levels of methylation in the cells. Constructs containing the lower two levels of expression (pNKR1707*Mse*Im, pNKR2213*Mse*Im did not result in full methylation of the cellular DNA, as judged by susceptibility of purified plasmid DNA from these cells to restriction by *Mse*I (1 µg plasmid DNA in 50 µl volume, 20 units *Mse*I, 1 hour at 37°C).

2) Construction of a library of randomly-mutagenized constitutive promoters by error-prone PCR

To find an intermediate level promoter construct for the *Mse*I methylase between that of pNKR2213MseIm and the apparently unstable pNKR1786MseIm, the constitutive promoter region was subjected to random PCR mutagenesis and selection. The mutagenesis protocol employed high levels of Taq DNA polymerase (5 units/100 µl reaction volume), unequal dNTP pools (1.2 mM dCTP and TTP; 0.2 mM dATP and

dGTP), high levels of $MgCl_2$ (7 mM), presence of $MnCl_2$ (0.5 mM), 2 ng of the pNKR1707MseIm per 100 µl volume and high PCR cycle numbers (35). The primers flanked the *Mse*I methylase gene at the *Age*I and *Bam*HI restriction sites respectively 5'-GCGATACAGACCGGTTCAGACAGGATAAAG-3' (SEQ ID NO:17) and 5'-GGTCGGATCCGGCGATACAGCGAG-3' (SEQ ID NO:18).

After PCR, the mutated promoter copies were restricted by *Age*I and *Bam*HI, gel purified with a Qiagen gel purification kit, and ligated into a *Age*I—*Bam*HI restricted pNKRMseIm construct that had been purified away from its endogenous constitutive promoter. Following electroporation into competent ER2688 cells, 20,000 colonies were achieved. These colonies were pooled and the plasmids were purified using a Qiagen purification protocol. This constituted a library of randomly mutagenized constitutive promoters, upstream of the *Mse*I methylase gene.

3) Selection of clones yielding plasmids resistant to *Mse*I restriction

To select for plasmids possessing a mutated constitutive promoter resulting in a stable, high level of methylation, 5 µg of the plasmid library was challenged by *Mse*I restriction (5 µg DNA, 50 units *Mse*I for 4 hrs at 37°C, followed by a 20 min incubation at 65°C to inactivate the *Mse*I restriction endonuclease. A portion of the challenged pool (250 ng) was

transformed into calcium-competent ER2688 cells and plated on supplemented LB agar plates and grown overnight at 37°C. This resulted in 63 colonies.

Six of these 63 colonies were randomly selected for further individual examination; after overnight growth in 10 ml supplemented LB medium, plasmid DNA was purified using a Qiagen Qia-prep spin miniprep protocol. When 100 ng of the purified plasmid DNA was challenged with 20 units of *Mse*I for 30 minutes at 37°C, all 6 were found to be fully restricted, indicating an inadequate level of methylation.

The remaining 57 colonies were pooled and a plasmid purification was done using a Qiagen plasmid purification protocol. From this plasmid pool, 50 ng was subjected to a longer (overnight) 50 unit *Mse*I challenge, followed by a 20 min incubation at 65°C to inactivate the *Mse*I restriction endonuclease. A portion of the challenged pool (4 ng) was transformed into calcium-competent ER2688 cells, plated on supplemented LB agar plates and grown overnight at 37°C. This resulted in 13 colonies.

Nine of these 13 colonies were randomly selected for further individual examination; after overnight growth and plasmid purification as previously described, 7 of the 9 were found to be fully methylated when 1 µg plasmid DNA was incubated with 50 units *Mse*I in a 50 µl reaction volume overnight at 37°C.

To further establish the level of methylation present in the cells, the 7 colonies were harvested for plasmid purification during the *logarithmic* phase of culture growth (cells were harvested 4 hours at 37°C after a 1:100 dilution of an overnight culture into fresh supplemented LB growth medium). Such cells would be expected to be replicating their DNA at such a rate that methylation by an expressed *Mse*I methylase might be unable to achieve complete methylation. Plasmid DNA was purified from these logarithmically growing cultures using Qiagen Qia-prep purification protocols and 0.5 ug of this plasmid DNA was incubated overnight at 37°C with 50 units *Mse*I. Using this more difficult methylation standard, 3 of the 7 colonies were fully protected (methylated) and resistant to restriction.

The three clones (#4, #9 and #10) resulting in a stable and full level of *Mse*I methylation had their promoter regions examined by mapping with *Age*I and *Bam*HI, and sequencing using a primer with an annealing position upstream of the promoter region. (5'-GGATCTTCCAGTGGTGCATGAACG-3' (SEQ ID NO:19). Two of the 3 clones (#9 and #10) were identical; thus the two step selection process described resulted in finding two independent promoters that yield a stable, full level of *Mse*I methylation.

Unexpectedly both promoter #4 and promoter #9/#10 were not mutagenized constitutive promoters as had been

the experimental design, but instead were *Age*I-*Bam*HI *E. coli* sequences that must have originated from the low level of *E. coli* DNA contamination present in the plasmid preparations.

The #4 promoter, by *Age*I/*Bam*HI, mapping appeared to be approximately 1000 bp in length; by sequencing, the first 438 bp were identical to *E. coli* K-12 MG1655 section 349 (Accession No. AE000459), base # 7813—8251. Upon examination of the sequence data, a *Bam*HI site was found at base #8814, which would yield the *Age*I-*Bam*HI *E. coli* fragment of 1002 bp. This *E. coli* sequence contains the 5' end of the yigW_2 orf and two predicted promoters, one of which is oriented in the same direction as the *Mse*I methylase (#8672-8704).

The #9/#10 promoter mapping appeared, by *Age*I/*Bam*HI, to be approximately 420 bp in length; by sequencing, the promoter was identical to *E. coli* K-12 MG1655 section 41 (Accession No. AE000151), base # 2511-2998. This defines a 488 bp *Age*I-*Bam*HI *E. coli* fragment that contains the 5' end of the *cof* orf and two predicted promoters oriented in the same direction as the *Mse*I methylase at positions #2605-2632 and #2714-2742. This #9/#10 sequence was used for further work.

4) Further optimization of *Mse*I methylase expression

Using the strategy described above, a level of *Mse*I methyltransferase expression which allowed expression of the *Mse*I endonuclease in plasmid pVR-25 was achieved. Unexpectedly, while the ER2566 host carrying the optimized *Mse*I methylase (#9 above) and the *Mse*I endonuclease in plasmid pVR-25 expressed *Mse*I endonuclease when first transformed and grown, the *Mse*I was not stably maintained when this construct was stored in glycerol at -70°C.

The *Mse*I methylase construct was further modified to achieve greater *Mse*I modification of the host. As described above, the attempts to place the *Mse*I methylase under the highest two levels of constitutive expression failed, presumably due to instability from high levels of methylation in the cells. To achieve a maximum tolerated level of methylation, a new M.*Mse*I expression plasmid, pVR-26, was constructed. pVR-26 was constructed by inserting a second promoter, derived as described in (3) above (see Table 1). This was done by cutting out a 1.244-kb DNA fragment containing the M. *Mse*I coding region (*mseIM* gene) and upstream promoter from plasmid pNKR1707mseIM-9 (digested with *Pme*I and *Mfe*I) and inserting it just downstream of the P_{lacUV5} promoter in vector pNEB193 (New England Biolabs, Inc., Beverly, MA) cut with *Eco*RI and *Hinc*II. Another *Mse*I methylase construct, pVR-27, was made by deleting a 0.379-kb *Pme*I-*Afl*III fragment containing the P_{lacUV5} promoter and

TABLE 1. Summary of plasmids and *Escherichia coli* hosts used for optimizing recombinant *Mse*I production

Strain	MseRM1	MseRM2	MseRM3	MseRM4	MseRM5	MseRM6
Host	ER2566	ER2566	ER2833	ER2833	ER2566	ER2833
pVR-25 PVR-24 with *mseIR*. At high temperature (42°C) without IPTG, the antisense promoter is active, while P_{T7} is repressed by LacI. At 30°C and with IPTG expression occurs from P_{T7}. Cm^R, ~10-15 copies/cell	+	+	+	+	+	+
pNK1707mseIM-9 pNK1707 with a *mseIM* gene and a fragment of *E.coli* chromosomal DNA containing promoter. Ap^R, ~40-50 copies/cell	+		+			
pVR-26 pNEB193 with *mseIM* fragment and upstream promoter region (*Pme*I-*Mfe*I fragment) from pNK1707mseIM-9. P_{UV5} and O_{lac} are active. Ap^R, ~500 copies/cell.		+		+		

Strain	MseRM1	MseRM2	MseRM3	MseRM4	MseRM5	MseRM6
Host	ER2566	ER2566	ER2833	ER2833	ER2566	ER2833
pVR-27 pVR-26 with a *Pme*I-*Afl*III deletion. The P_{UV5} and O_{lac} are deleted. Ap^R, ~500 copies/cell					+	+
pCEF-8 pSYX20 (pSC101 origin) with T7 lysozyme gene cloned into *Sal*I site in oposite direction to the P_{Tet}. Kn^R, ~2-5 copies/cell	+	+	+	+	+	+
lacI^Q			+	+		+
LacI copies/cell	110-160	110-160	210-260	210-260	110-160	210-260
O_{lac} copies/cell	13-18	~500	13-18	~500	13-18	13-18
Drug resistace	Ap^R,Cm^R,Kn^R	Ap^R,Cm^R,Kn^R	Ap^R,Cm^R,Kn^R	Ap^R,Cm^R,Kn^R	Ap^R,Cm^R,Kn^R	Ap^R,Cm^R,Kn^R

lacI operator from pVR-26. The pVR-26 *mse*IM methylase expressing vector allowed the stable expression of *Mse*I endonuclease.

EXAMPLE V

Optimization of the *Mse*I restriction endonuclease expression

1) Expression vector construction

As known very well in the art, restriction endonucleases are cytotoxic proteins. Attempting to clone a toxic gene into a plasmid designed to facilitate high expression is, in many cases, extremely difficult. One especially preferred plasmid for expressing cytotoxic genes is pLT7K (Kong, et al., *Nucl. Acids Res.* 28:3216-3222 (2000)). This plasmid contains the segment encoding replicative function (ori), a gene encoding ß-lactamase, and a gene encoding kanamycin resistance which is flanked by restriction endonuclease sites suitable for cloning. The cI857 gene encodes a mutant form of the of the lambda bacteriophage repressor protein, which conditionally binds to DNA sequences (the CI operator) that overlap PL and PR (the lambda bacteriophage major leftward and rightward promoters, respectively). The lacI gene encodes a repressor protein, LacI, that conditionally binds a DNA sequence (the lac operator) which has been constructed to overlap PT7 (bacteriophage T7 RNA polymerase transcriptional promoter). Briefly, at high temperature (42°C) without IPTG, the

antisense promoter is active, while P_{T7} is repressed by LacI. At 30°C and with IPTG, expression occurs from P_{T7}.

To adapt the pLT7K for overexpression of *Mse*I restriction endonuclease gene, an NdeI restriction endonuclease site and ribosome binding site were introduced. Additionally, the colEI replicon was changed to the p15A replicon and copy number was decreased 3 times (from ~50 to ~15). To acomplish this, pLT7K was digested with *Acl*I and *Bam*HI. The resulting 1.2-kb fragment containing cI857, the lambda PL, Kn resistance gene and the T7 promoter was isolated from an agarose gel using Qiagen QIAquick Gel Purification Kit (Cat. No. 28704) and ligated into pACYC184-T7terΔPshAI vector that was previously digested with *Cla*I and *Bam*HI. The pACYC184-T7terΔPshAI is a *Psh*AI deletion derivative of pACYC184-T7ter. This construct was designated pVR-24 (Fig. 11).

The open reading frame (ORF) for the *mseIR* gene was amplified by PCR with a set of forward (5' AGACTCCCCCATAT GACCCACGAACCGACGGATG 3' (SEQ ID NO:20) and reverse (5' GGGTGGTCCCGCTAGCTATTAGTAGGGACCGGGG 3' (SEQ ID NO:21) primers, where the underlined bases show the positions of the NdeI cleavage site for the forward primer. PCR was performed using Vent® DNA polymerase, 1X ThermoPol buffer, 500 ng of *Micrococcus* species (NEB#446) chromosomal DNA as the template in a 100 µl PCR reaction, and primers. Twenty-five cycles were done consisting of 15

sec at 95°C, 60 sec at 68°C and 45 sec at 72°C. The resulting 700 bp PCR product was purified using a QiaQuick PCR purification protocol, treated with Klenow fragment, digested by *Nde*I , and purified once again using the QiaQuick PCR purification protocol.

The resultant 700-bp *Nde*I-Blunt end fragment, containing *Mse*I restriction endonuclease gene,was ligated into pVR-24 vector digested with *Nde*I and *Stu*I and ligation mixture was transformed into *E. coli* ER2502 cells, previously modified with the *Mse*I methylase gene construct pNKR1707*Mse*Im-9. Out of 18 individual transformants analyzed, three contained *mseIR* gene. After sequencing the DNA insert containing *Mse*I restriction endonuclease gene, one recombinant plasmid, pVR-25, was selected for producing the *Mse*I restriction endonuclease.

2) Strain Construction

To increase LacI repressor copy number in the host, the strain ER2833 (T7lacIq strain) was constructed as described in U.S. Application Serial No. _____.

3) Optimization of *Mse*I restriction endonuclease overexpression in *E. coli* combining different hosts and plasmids expressing different levels of *Mse*I methylase.

For optimization of *Mse*I restriction endonuclease overexpression in *E. coli*, the pVR-25 plasmid was transferred

into the expression strain ER2566/pCEF-8, which was pre-protected against *Mse*I endonuclease auto-digestion by carrying one of these *Mse*I methylase expressing plasmids (pNKR1707*Mse*Im-9, pCR-26 and pVR-27). ER2566/pCEF-8 is a host strain containing a chromosomal copy of the gene for T7 RNA polymerase under control of the inducible lac promoter and a pSYX20 based plasmid, pCEF-8, which specifies low levels of T7 lysozyme, a natural inhibitor of T7 RNA polymerase. For additional information, see Moffatt, B. A., and Studier, F. W., "T7 Lysozyme inhibits transrciption by T7 RNA polymerase," *Cell*, 49:221-227 (1987). In uninduced cells, lysozyme reduces the basal activity of the T7 RNA polymerase and increases the range of target genes that can be stably maintained in the expression host. In addition, another expression strain, ER2833/ pCEF-8 was used, which has an copy of lacIq gene on the F' episome.

Overall, six strains were used of *Mse*I restriction endonuclease expression studies in *E. coli* (Table 1). All strains contain pVR-25 plasmid, expressing *Mse*I restriction endonuclease, and pCEF-8 plasmid which encodes a T7 bacteriophage lysozyme gene. A variety of growth conditions were employed to grow transformed host cells to select for higher yields of *Mse*I restriction endonuclease. The preferred medium in optimization experiments was Luria-Bertani (supplemented with 1 gram glucose and 1 gram MgCL2 per liter; subsequently referred to as supplemented LB) media.

The growth conditions were as follows:

MseRM1: cells from an individual colony were grown in 0.5 liter of LB medium at 42°C for 8h, after which IPTG was added to 0.2 mM final concentration to induce the T7 RNA polymerase and cells were grown overnight (15 h) at 30°C. Antibiotics were added as needed: 30 µg of kanamycin per ml, 100 µg of ampicillin per ml, and 30 µg of chloramphenicol per ml. Finally, cultures were harvested by centrifugation and frozen at -20°C.

MseRM3: for each experiment, cells from an individual colony were grown in O.5 liter of LB medium at 30°C overnight (17 h), after which IPTG was added to 0.2 mM final concentration to induce T7 RNA polymerase and cells continued to grow for 4 h. Antibiotics were added as needed: 30 µg of kanamycin per ml, 100 µg of ampicillin per ml, and 30 µg of chloramphenicol per ml. Finally, cultures were harvested by centrifugation and frozen at -20°C.

MseRM4, MseRM5 and MseRM6: bacterial cultures were kept as frozen stock solutions at 70°C in 50% glycerol. Cultures used for seed inoculation were streaked onto LB medium plates containing the appropriate antibiotics to obtain single colonies. An individual colony was resuspended in 1 ml of LB medium and inoculated into a 1000-ml flasks containing 500 ml of LB medium supplemented with 30 µg of kanamycin/ml 100 µg of ampicillin/ml, and 30 µg of

chloramphenicol/ml. Cells were grown overnight (16 h) in a shaking incubator at 37°C and 250 rpm. Thereafter, IPTG was added to a final concentration of 0.2 mM. Cells were cultivated for another 4 h and then were harvested by centrifugation at 8,000 g for 5 min at 4°C and frozen at -20°C.

Two preferred restriction endonuclease assays for identifying high-level expression clones were used.

Sonication method: induced cultures (500 ml) were harvested and resuspended in 20 ml sonication buffer containing 10 mM Tris.HCI (pH 7.5) and 1 mM EDTA. Cells were sonicated on ice by four 30 second blasts with a macro-tip probe. A portion of the crude extract was added to lambda DNA (1µl) in NEBuffer 2 buffer (50 µl) and incubated for 1 hour at 37°C. DNA was fractionated by 0.8% gel electrophoresis and visualized by EtBr staining.

EXPRESS method: one ml of an overnight or induced culture (10 - 500 ml) was harvested and resuspended in 0.2 ml buffer containing 50 mM TRIS-HCl, pH 7.5 and 25% (vol/vol) sucrose and mixed until the solution was homogenous. 11 µl of 200 mM EDTA, pH 8.0 plus 200 µl of freshly-prepared 10 mg/ml lysozyme in 0.25M Tris-HCl (pH 8.0) were added and the solution was incubated on ice for 5 min. 11.5 µl of 1 M $MgCl_2$ and 24.2 µl of 5% (vol/vol) Brij-58 were then added. The solution was gently mixed and incubated in room temperature for 15 min. After incubation the crude cell lysate

was centrifuged at maximum speed in a microcentrifuge for 15 min at 4°C. The supernatant was pipetted off into a new eppendorf tube and stored on ice until needed. Lambda DNA substrate (1.0 μg) was digested in *Mse*I reaction buffer buffer (NEBuffer 2) with serial dilutions of cell extract for 1 hour at 37°C degree. DNA was fractionated by electrophoresis and visualized by EtdBr staining. Activity was determined by the presence of the appropriate size bands associated with a *Mse*I digestion of lambda DNA.

The results of optimization of *Mse*I restriction endonuclease expression are summarized in Table 2.

MseRM1 strain gave a variable yield of *Mse*I restriction endonuclease (0.08-0.5x10^6 U/g wet cells). Cells grew slowly and the lag time was exceptionally long.

To enhance the stability and reproducibility of lac-based recombinant expression systems, the new host strain ER2833 (U.S. Application Serial No. _____) was constructed, which has an copy of lacI^q gene on the F′ episome. Indeed, the expression stability and plasmid maintenance in the lacI^q host (MseRM3) was greatly enhanced: the yield of *Mse*I restriction endonuclease was 0.5-1.4x10^6 U/g wet cells. The *Mse*I restriction endonuclease purified from this strain (see Example VI) was substantially free of non-specific endonuclease and exonuclease and the final yield was

Table 2. Summary of optimization of MseI restriction endonuclease expression in *E.coli*

Strain	Induction Conditions	Yield (U/g)	Comments
MseRM1	42°C 8 h (~40 Klett), shift to 30°C overnight	0.08-0.5x10^6	Difficult to repeat results. No activity from frozen culture
MseRM2			ER2566/pVR-26 grew very slowly, impossible to make competent cells
MseRM3	30°C 37°C overnight (~100 Klett), IPTG (0.2 mM) 4h	0.5-1.4x10^6	No activity from frozen culture. The enzyme prep gave ~150,000 U/g
MseRM4	37°C overnight (~100 Klett), shift to 30°C + IPTG (0.2 mM) 4h	3.3-8.6x10^6	This strain gave stable results from frozen culture, high *Mse*I yield
MseRM5	37°C overnight (~100 Klett), shift to 30°C + IPTG (0.2 mM) 4h	1.5-3.4x10^6	This strain gave stable results from frozen culture, but has less *Mse*I yield than MseRM4
MseRM6	37°C overnight (~100 Klett), shift to 30°C + IPTG (0.2 mM) 4h	3.3-3.8x10^6	This strain gave stable results from frozen culture, but has less *Mse*I yield than MseRM4

[a]Resuspend an overnight colony (plated on 42°C) in 1ml LB, then add 0.1 ml of resuspended colony into the flask containing 500 ml of LB + antibiotic. Grow as described in EXAMPLE IV.

~150,000U/g. It is about 100 times greater yield than from native *Micrococcus* species (NEB#446).

Unfortunately, the MseRM3 strain showed no *Mse*I restriction endonuclease activity after the strain was stored at -70 °C and revived. To solve this problem, the *Mse*I methylase expression level was increased by constructing pVR-26 and pVR-27 plasmids (Example IV above). These strains (MseRM4, MseRM5 and MseRM6) gave high *Mse*I restriction endonuclease yield from even after storing the strain at -70 °C and one strain, MseRM4 (NEB#1284; New England Biolabs, Inc., Beverly, MA) was used for scale-up in the 100 L production fermentor (see Example VI). The yield of *Mse*I restriction endonuclease from this larger scale fermentation was 0.5×10^6 U/g wet cells.

EXAMPLE VI

Production of the recombinant *Mse*I restriction endonuclease

The *Mse*I restriction enzyme was produced from recombinant *E. coli* strain NEB#1284 propagated to late-log phase in a 100-liter fermenter. A sample of these cells was deposited under the terms and conditions of the Budapest Treaty with the American Type Culture Collection on August 28, 2000 and received ATCC Accession No. PTA-2421.

A) Cell Growth

The transformed *E. coli* host, NEB#1284 containing the recombinant *Mse*I restriction endonuclease clone was stored as a frozen stock solution at -70°C in 50% glycerol. Cultures used for seed inoculation were streaked onto LB agar plates containing ampicillin (100µg/ml), chloramphenicol (30µg/ml) and kanamycin (50µg/ml) and incubated overnight at 37°C to obtain single colonies. Several colonies were used to inoculate 10 ml LB medium supplemented with 30 µg of kanamycin/ml 100 µg of ampicillin/ml, and 30 µg of chloramphenicol/ml. Cells were grown for 3 hrs in a shaking incubator at 37°C and 250 rpm and then at 30°C for an additional 3.5 hours (to avoid overgrowing the culture). The final corrected Klett of this culture was 122 or mid-log. This culture was used to inoculate 100-liter of LB supplemented with 30 µg of kanamycin/ml 100 µg of ampicillin/ml, and 30 µg of chloramphenicol/ml. The fermentation was run for 18 hours at 30°C with aeration of 2 SCFM (standard cubic feet per minute) and an agitation rate of 200 rpm. The final corrected Klett was 313. From this fermentation 331 grams of cells (wet weight) were harvested by continuous flow centrifugation and cells were stored at -70°C. A crude extract was made from 1 g of cells and the enzyme activity was estimated, using the method described above (see Example V). The yield of *Mse*I restriction endonuclease in crude extract was 500,000# U/g, which is about 100 times more than in crude extract of *Micrococcus species* (NEB#446).

B) Purification of the *Mse*I restriction endonuclease from NEB# 1284

All the following procedures were performed either on ice or at 4°C. 330 grams of cells were suspended in 990 ml Buffer A (0.15 M NaCl, 10mM Tris pH 7.5, 10 mM BME, 1 mM EDTA and 5% (v/v) glycerol and were broken by 4 passes at psig12K through a Gaulin Press to an O.D. of 0.56. The 1150 ml supernatant was PEG precipitated by adding PEG 6000 to 7.5% and NaCl to 0.5 M and then incubated for 50 minutes at 4°C. The PEG slurry was centrifuged at 12K for 30 minutes at 4°C. The 580 ml of supernatant was diluted to 0.1M NaCl with Buffer A without NaCl and loaded onto a 430 ml Heparin Hyper D column equilibrated with Buffer A. The column was washed with 1200 ml Buffer A and then a 4000 ml linear gradient from 0.1 M NaCl to 1.0 M NaCl was applied. The restriction enzyme activity eluted at 0.25-0.35M NaCl and was pooled. The Heparin Hyper D pool was diluted to 0.1M NaCl with Buffer A without NaCl and loaded onto an 88 ml PEI column equilibrated with buffer A. The column was washed with 100 mls Buffer A and then a 1000 ml linear gradient from 0.1M to 1.7M NaCl was applied. The restriction enzyme activity eluted at 0.7 - 0.9M NaCl and was dialyzed against Buffer C (50 mM NaCl, 15 mM Tris pH 7.5, 10 mM BME, 0.1 mM EDTA and 5% (v/v) glycerol) overnight and loaded onto a 20 ml Source Q column equilibrated with Buffer C. The column was washed with 40 ml Buffer C and a 400 ml linear gradient from

0.05M NaCl to 1.0M NaCl was applied. The restriction enzyme activity eluted at 0.25M -0.35M NaCl and was pooled. The Source Q pool was dialyzed against Buffer D (10 mM KPO4 pH7.0, 0.075M NaCl, 10 mM BME, 0.1mM EDTA, 5% (v/v) glycerol) and loaded onto a 20 ml Heparin TSK column equilibrated with Buffer D. The column was washed with 40 ml Buffer D and a 400 ml linear gradient from 0.075 M to 1 M NaCl in buffer D was applied. The restriction enzyme activity eluted at 0.3M -0.4M NaCl and was pooled. BSA was added to a final concentration of 100 µg/ml. The pool was dialyzed to Storage Buffer (20 mM Tris pH 7.5, 0.1M EDTA, 1mM DTT, 50 mM NaCl, 50% (v/v) glycerol, 200µg/ml BSA) overnight. This purification scheme yielded 26,000,000 units of *Mse*I restriction endonuclease. The *Mse*I restriction endonuclease obtained from this purification was substantially free of non-specific endonuclease and exonuclease.

The purity of the *Mse*I restriction endonuclease preparation was checked by looking at the following criteria:

1. Ligation: After a 5-fold overdigestion of lambda DNA, greater than 95% of the DNA fragments produced were ligated with T4 DNA Ligase (at a 5′ termini concentration of 1-2µM at 16°C. Of these ligated fragments, 95% were able to be recut.

2. Prolonged digestion: After incubating a 50µl reaction containing 1 µg of lambda and 100 units of enzyme

for 16 hours, the same banding pattern of DNA bands was produced as a reaction performed in one hour with one unit of enzyme.

3. Exonuclease Activity: After incubation of 100 units of enzyme for 4 hours at 37°C in a 50 µl reaction containing 1 µg sonicated ^{3}H DNA (10^{5} cpm/µg) less than 0.4% radioactivity was released.

All tests were performed in the following reaction buffer: NEBuffer 2 (50 mM NaCl, 10 mM MgCl2, 10 mM Tris-HCl, 1 mM DTT, (pH 7.9 at 25°C, supplemented with 100 µg/ml BSA. Unit determination: Lambda DNA substrate (1.0 µg) was digested in 50 µl 1X *Mse*I reaction buffer (NEBuffer 2) with serial dilutions of *Mse*I endonuclease for 1 hour at 37°C. DNA was fractionated by electrophoresis and visualized by EtdBr staining. Activity was determined by the presence of the appropriate size bands associated with a *Mse*I digestion of lambda DNA. One unit of restriction endonuclease activity is defined as the amount of enzyme required to completely digest 1 µg of substrate DNA in a total reaction volume of 50 µl in one hour using the NEBuffer specified.